Response to the Reviews and revised manuscript with marked changes for 'Observed groundwater temperature response to recent climate change' by K. Menberg et al.

Reply to review by H. Kooi:

General comments

Comment #1: The work presented in this manuscript involves a case(s) study of the impact of climate warming (atmospheric temperature rise) on groundwater temperatures at depths of about 10-30 m. This topic is relevant to the hydrological community. The authors draw particular attention to the potential relevancy for stream/river temperatures and (associated) groundwater dependent ecosystems.

> **Reply #1:** This statement accurately captures the scope of this paper.

Comment #2: Novel and/or particularly interesting aspects of the work are the rather long time series of observational data on (pumped) groundwater temperature and the 'regime shift analysis' approach. The advantage of the latter is that it allows to establish in an elegant way a relation between local (diffused) 'shifts' in groundwater temperature and climatic regime shifts that can be recognized over very large spatial domains, even up to the global scale. It is interesting to see that it apparently works. However, applicability of the method to the 'diffused signals', which inherently do not include 'abrupt shifts' needs to be justified and findings need to be interpreted more carefully to avoid misinterpretations of air-ground temperature coupling (SAT-GST).

> **Reply #2:** We agree that the applicability of the analysis of 'abrupt shifts' in groundwater temperatures is certainly limited due to the diffusive heat transport in the subsurface. Abrupt shifts in the thermal signal at the surface will diminish within the subsurface and actually disappear in a certain depth, appearing rather as diffused signals. However, in the quite shallow GWT time series evaluated here, the breakpoints observed in the long term mean GWT are statistically significant, though exhibiting lower *p*-values than in the atmospheric time series. Additional information on this issue and the statistical requirements of the regime shift analysis are addressed in detail according to the specific comment #22. The interpretation of the observed regime shift in GWT regarding the damping of the thermal signals and the adjacent

discussion of the air-ground temperature coupling (SAT-GST) was thoroughly revised according to these concerns and the comments #17, #20 and #22.

Comment #3: Where virtually all geothermal climate studies use borehole temperature logging data, here temperatures obtained by well pumping are used. This aspect (value of this type of data) deserve to be elaborated more comprehensively and better, because results basically confirm what is already known about propagation of surface temperature signals into the subsurface or air-ground temperature coupling.

Reply #3: As also stated correctly in comment #12, the novelty of our study is the type of data used for the analysis of air-ground temperature coupling. The measured time series of groundwater temperature and the applied methods enable an evaluation of the temperature coupling over the last decades with a temporal resolution of a few years. Palaeoclimatic studies using deep borehole temperature profiles can track ground temperatures over hundreds of years before present, but with a much lower temporal resolution in the order of decades. A respective statement was included in the manuscript to highlight these distinctions (p. 4, lines 12-16).

Due to general knowledge about the propagation of thermal signals in the subsurface, the observed influence of atmospheric temperature signals on groundwater temperature might indeed be anticipated. However, according to our knowledge this is the first study that examines this short term coupling by statistical analysis of actual measured data and analytical modeling. Details about the newly derived implications of this short term coupling for groundwater dependent ecosystems are addressed with the associated comments #12 and #25.

Comment #4: The conclusions provided in the final conclusions section are generally sound. However, the forward modelling with the analytical solution, its results, and the discussion thereof is inadequate in several respects and causes unnecessary confusion as will be explained below. These aspects can be remedied fairly easily.

Reply #4: Thank you for raising these concerns. They are addressed on a point by point basis below according to the comments #9, #10, #17 and #21-25.

Specific comments

Comment #5: GWT data. Compared with temperature logging in standing water in a well bore, interpretation of temperatures obtained via well pumping is subject to a large number of unknown influences. Firstly, the water represents a mix of water entering the well bore along the whole! vertical of the well screen and hence different depths. The inflow can be fairly even, but can also have a dominant inflow near the base or the top of the screen depending on aquifer heterogeneity, backfill and screen clogging processes. This distribution would even depend on the pumping rate/ induced water level drop in the well bore during pumping. This is of fundamental importance for the Hardtwald wells which have very long screens.

> **Reply #5:** Thank you for noting the influence of the well screen depth and how heterogeneities can influence the vertical well capture zone. These comments are addressed in more detail below in the replies to comments #8-10 and #21. Based on your concerns, we included a statement in the methods that describes why we now consider a range of depths in the analytical model (p. 10, lines 11-15).

Comment #6: Secondly, depending on the heterogeneity structure of the aquifer around the well bore, there could in principle even be a relatively large groundwater contribution from below the bottom of the well screen or above its top.

> **Reply #6**: We employed this information to present a plausible reason for the measured Sinthern GWT not adhering to the predicted GWT trends (p. 14, line 30 – p.15, line 4).

Comment #7: Thirdly, the water flows upward from the pump through the pumping hose/tube which exchanges heat with the air in the well bore, air outside, and, a notorious factor is the impact of direct solar insolation of the hose which can heat up even fairly fast flowing water very quickly. The heat exchange very much depends on factors such as pumping rate and tube size, length and material, and ambient air temperature during sampling (season). A constant outflow temperature for a constant pumping rate does not guarantee the temperature is representative of the groundwater temperature. A simple check for seasonal trends in the data would be a minimum (for the mean parameters used for Dansweiler, for instance, at the depth of the screen (20 m), numerical modelling shows a seasonal GST change between 0 and 20°C corresponds to a groundwater temperature fluctuation of about 0.02°C). Much larger fluctuations point at ‘contamination effects ’such as those mentioned above. The observed inter-annual fluctuations of several tenths of a degree for the Dansweiler well already indicate that such influences are significant. Furthermore, has the whole procedure (protocol and instrumentation) used for sampling been exactly the same over the 40 years? This is an additional potential cause of fluctuation or systematic changes. Such uncertainties should be acknowledged/considered when using the data.

Reply #7: We agree that there are several potential reasons for the rather large fluctuations in the inter-annual GWT time series, which are not properly addressed in the manuscript so far. We made the according changes to reflect the uncertainties of this type of data (p. 5, line 29 – p. 6, line 5) and the subsequent implications for the interpretation (p. 11, lines 25-30). The major concerns are briefly addressed here.

Due to the length of the time series of up to 30 years, detailed information about the pumping rate during sampling, the used equipment or ambient air temperatures are not available for the whole time period. The GWT time series from all wells exhibit certain seasonality with temperature measurements in summer and autumn being on average 0.4°C higher than in winter in Dansweiler, and 0.5°C in the Hardtwald wells. These values are considerably higher than the predicted fluctuations of 0.02°C from the seasonal SAT changes, which means that there is most likely an impact of ambient air temperature on the groundwater in the hose during sampling. Yet, it should be noted that the accuracy of the measurements is ±0.1°C, which is close to the observed variations. Consequently, the sentence stating that the measured temperature is representative for the upper aquifer will be deleted. However, years with only one measurement are scarce, so that annual mean GWT is rarely influenced by these seasonal fluctuations. Individual outliers in the time series caused by such variations or due to other reasons are accounted for in the statistical analysis and do not influence the results of the regime shift analysis due to the sequential approach (see comment #22 on statistical requirements below).

The measurements were performed by the water authorities within the groundwater quality observation program. The measurement protocol, which is standardized by the environmental state agencies to assure good quality of the obtained data, has undergone no significant changes in the last decades (p. 5, lines 25-27). The instrumentation will have certainly changed within this rather long period, which is now addressed as a possible source for GWT fluctuations in the manuscript (p. 6; lines 3-6).

Comment #8: Modelling and its interpretation. Given the general groundwater flow behaviors for pumped wells mentioned above, it is conceptually inappropriate to compare the observed temperature time series with a model-generated time series for the depth corresponding to the water table. Most logical would be to generate a time series for the mean temperature (integral divided by length) along the depth of the screen. Even depth-weighted integrals could be considered in a sensitivity analysis for uneven inflow into the well. The 'cone of depression' (p. 3654, line 8) is not a concept which would justify the adopted water table depth approach.

Reply #8: We agree that our original approach was not ideal, and we thank you for raising this concern. We solved out the integral of the analytical solution, and the resultant equation was several lines long. Given the approximate nature of this study, we think it is best to rather simply use the midpoint of the screen depths for our 'best guess' results (p. 14, lines 17-19). A quick glance at Figure 4 indicates that the temperature results corresponding to halfway down the well screen are approximately halfway between the upper and lower limits of the temperature envelope. These upper and lower limits represent results for the top and bottom of the well screen, respectively (p. 14, lines 10-17).

Comment #9: For Dansweiler and Sinthern a single depth of about 20 or 21 m may be appropriate because of the short well screen. For the Hardtwald wells an integral approach is crucial; the water table depth (6 m) definitely is way too shallow to generate a meaningful time series. This most likely accounts for the inferred offset between ΔSAT and ΔGST for these wells, which therefore seems an artifact. The text of sections 2.3 and 3.2 should be modified accordingly.

Reply #9: We now consider depths that range from the top of the well screen to the bottom (p. 10, lines 11-15; p. 14, lines 10-19; p. 15, lines 22-26). As you suggest, this change to the manuscript eliminates any suggestion of an offset between ΔSAT and ΔGST. We have removed such statements and are pleased with how the manuscript has been improved based on your suggestion.

Comment #10: Presently, the predictive uncertainty of the model is captured in Figure 4 in the 'predicted GWT range'. However, this is due to uncertainty in thermal parameters only. The uncertainty caused by the screen length in combination with unknown inflow distribution is way larger. Point depths ranging between 15 and 25 m may be reasonable estimates for this uncertainty (or specified uneven inflow distributions with the integral approach).

Reply #10: Good suggestion. We now include uncertainties in the predicted envelope due to thermal properties AND depths. The maximum limit of the temperature envelope was obtained using the depth to the top of the well screen and the highest diffusivity and lowest heat capacity (heat capacity is in the U term in the solution), whereas the minimum limit of the temperature envelope was obtained using the depth to the bottom of the well screen, the lowest diffusivity, and the highest heat capacity (p. 14, lines 10-21).

Comment #11: Table 3 lists ranges of thermal parameter values. However, the combination of heat capacity and thermal diffusivity values in not clear from the way they are presented. Probably a small bulk heat capacity would correspond to a large diffusivity, otherwise thermal conductivities seem unrealistic. This should be clarified.

Reply #11: In order to clarify this issue, we included the assumed literature values of the thermal conductivity of the saturated and unsaturated zone in Table 3 (p. 27). The given ranges in thermal conductivity and heat capacity account for varying water saturation of the porous media and for variation in the composition of the sedimentary material, i.e. different contents of gravel, sand, etc. Due to the interaction of these variations the small bulk heat capacities in Table 3 do not necessarily correspond to the larger diffusivities. The corresponding paragraph in the manuscript was changed accordingly (p. 10, lines 1-10).

Other comments/corrections

Comment #12: p. 3638: line 35: Rather vague and in my opinion incorrect statement. In what sense are the implications of climate change for groundwater temperatures not comprehensively understood? The present study certainly does not add to or require changes in present understanding. What is shown (with corrections suggested) was predicable on forehand. What is new here is that it is shown that long temperature time series obtained from pumping wells can also be valuable to document and study climate impacts, in spite of its more ‘contaminated’ and vertically integrated signature.

Reply #12: We do not necessarily agree. There is, in fact, a poor understanding of how shallow groundwater temperature rise may impact ecologically important aquifers and rivers. This is manifested by the plethora of surface water temperature papers that consider stream warming due to climate change but do not consider the potential of groundwater temperature rise to influence stream warming. Surface water hydrologists are one of our target audiences for this paper, and this is partly why we have included so much ecohydrological content. In particular, our study shows that groundwater may warm rapidly and drastically in response to climate change, which are surprising results to some (although not necessarily to anyone who understands subsurface heat transport). The matter of predictability of our results and the novelty of the used data and methods is already described in the reply to comment #3.

Comment #13: line 10: Abrupt changes in groundwater temperature? Violates heat transport behavior.

Reply #13: We agree that this was not a good descriptor of the diffuse signal. We meant abrupt increases in terms of breakpoints in the long term mean. This was confusing; hence 'abrupt' has been removed from this sentence in the abstract (p.1, lines 23-24) as well as a related sentence in the conclusions (p.17, line 19).

Comment #14: p. 3642: lines 16/17: Variations of water table of 6m (and mean water table 6m below land surface beg for some explanation. Relevancy for the present study, and the magnitude in relation with the recharge of about 220 mm/yr. Is there a pumping station nearby? Irrigation extraction by farmers? How can this be consistent with a steady vertical advective heat flow (U) in the model?

Reply #14: The variation in the depth of the water table stated here is the total variation, i.e. the mean depth is 7 m, with a maximum depth of 10 m and a minimum depth of 4 m, which occurred in individual years during observation time. We rephrased the variation to 'maximum variations of ±3m' to make this clear (p. 5, lines 10-12).

Variations in the depth of water table are relevant for the interpretation of the time lags of the shifts between atmospheric and groundwater temperature. In particular, a trend in the depth of the water table would impair the comparison of time lags of different shifts in one well. However, there is no obvious long term trend in the water level of the four observation wells over the last decades. The wells in the Hardtwald are located near a pumping station, which is likely to influence the water level. According statements were included in the manuscript (p. 5, lines 11-13).

Time-series for the annual groundwater recharge were unfortunately not available for the whole observation time. To account for the long-term variability an uncertainty of ±20% was assigned to the annual recharge values, which were also applied to the vertical advective heat flow U in the model (p. 10, lines 16-20).

Comment #15: p. 3646: section 2.3: Would be good to also explicitly state the model assumes (a) uniform and steady vertical groundwater flow over a depth range deeper than the well depths and (b) recharge temperature equals the average annual surface temperature.

These assumptions, together with assumed heterogeneity of thermal properties for a variably saturated system, merit discussion in later sections in relation to conclusions drawn from the modelling.

> **Reply #15:** These are certainly assumptions associated with the governing conduction advection equation. We noted these, along with others, in the methods section (p. 8, lines 4-8). We also included a new paragraph in the discussion that explains the shortcomings of this approach (p. 16, lines 6-21).

Comment #16: p. 3648: Equation (6): For sake of completeness mention that the contribution of each summation term only applies for t >= t_i. Otherwise unwarranted cooling is calculated before the relevant step change in surface temperature. U_z is not defined and appears to equal U.

> **Reply #16:** This is precisely what the Heaviside function indicates (i.e., the Heaviside function turns on and stays on when the value inside the Heaviside function is positive). Nonetheless, such a statement was added (p. 9, lines 13-14). U_z should have been Uz as can be shown from an analysis of the units within the exponential term (Eq. 6).

Comment #17: p. 3647, line 17-25: This is inappropriate reasoning. In the model initial GWT and hence GST are set equal to observed GWT. Potential offsets in SAT GST due to surface conditions in the real world system are subsequently of no consequence for the imposed step changes in annual GST (unless a step change in surface conditions (e.g. vegetation or snow regime) occurred at the same time, which is not the case).

> **Reply #17:** As other studies have shown (Kurylyk et al., 2013, HESS; Mellander et al, 2007, Clim. Change), a shift in SAT can produce a shift in snowpack conditions and/or deciduous vegetation, which produces decadal GST changes that do not necessarily follow SAT changes. However, this text has been removed from the manuscript based on the modifications we have made to the depths utilized in the analytical solution. In the approach of the modified manuscript we have set ΔGST = ΔSAT for all wells (p. 9, lines 25-31).

Comment #18: p. 3650, lines 3-5: This is a vague statement, in particular the ‘up to 30 m’ and ‘significant’. It can be readily shown that for the well sites studied here variations due to inter-annual fluctuations of GST (or SAT) are much smaller than those observed in the data.

Analytical solutions to quantify the damping of periodic GST fluctuations with depth (also with advection influences) can be used to show this. Or numerical solutions can be used.

> **Reply #18:** We agree that this was poorly worded. Stallman's (1965) equation could be used to demonstrate that intra-annual fluctuations should be completely damped at this depth even under high recharge rates. The sentence was removed.

Comment #19: p. 3650: If the accuracy of each shift is +/- 1 year, then the accuracy of the difference between two shifts (lag) is less accurate than that.

> **Reply #19:** We agree. As the accuracy of ±1 year applies to the shift in SAT and to the succeeding shift in GWT, the overall accuracy of the difference is ±2 years. The values in Table 4 was changed accordingly (p. 28).

Comment #20: lines 18-20: What is the relevancy of this statement?

> **Reply #20:** The point is that local SAT will not necessarily follow global SAT or even regional SAT changes. We included the adjective 'local' and change 'yet' to 'furthermore' to clarify the purpose of this statement (p. 12, lines 12-15). This helps explain why the timing of regime shifts at different spatial scales may not completely overlap.

Comment #21: lines 23-25: This is not substantiated and not evident. The depth of the well screens may be more important (can be evaluated via sensitivity analysis).

> **Reply #21:** This sentence was deleted. The sensitivity of the GWT response to the well screen depth is now considered in the results 'envelope' in the analytical solution figure (p. 14, lines 10-19; Fig. 4, p. 33).

Comment #22: p. 3651, lines 8-12: Indeed. Due to this slow and 'smoothed' response in the subsurface I would expect the regime shift method is NOT suited to determine the proper amplitude of the GWT and hence the GST shift, and overestimate its timing. The inferred amplitude step change of the diffused signal would depend on the length of the stable regime. The inferred amplitude can therefore NOT be used to draw conclusions regarding damping of GST change relative to SAT change. Aren't there statistical requirements of time series for regime shift analysis? And do diffused signals meet these requirements? The discussion of lines 25-28 seems inappropriate in this light.

> **Reply #22:** We agree that the regime shift method can underestimate the groundwater rise in response to climate change (unless equilibrium has been met due to long stable

regime). That is why, unlike Figura et al. (2011, GRL), we include the process-based modeling which demonstrates that GWT will eventually rise (with the same magnitude) in response to the regime shift in SAT or GST if the new climate regime shift lasted indefinitely. We agree that no conclusion regarding ΔGST damping in comparison to ΔSAT can be made. Such statements were removed. We now explicitly state that in the absence of snowpack evolution or land cover changes, ΔGST should follow ΔSAT, and ΔGWT should in turn follow ΔGST if given enough time (p. 15, line 27 – p. 16, line 5; p. 17, lines 27-29).

There are requirements regarding the length of a time series in order to detect the regime shifts within a certain level of confidence depending on the assigned length of a stable regime. With an assigned regime length of 10 years, as in our study, up to 3 statistical significant regime shifts could theoretically be identified in a time series of 40 years. In principle the regime shift analysis resembles the fitting of a step function to a time series, though with a limited number in the degrees of freedom. Not only is the length of the identified regimes restricted by a minimum value (cut-off length), also the difference in the long term means between two subsequent has to be statistically significant according to a student's t-test. In this test also the variance of the input data is considered, which means that the discussed large fluctuations in GWT are accounted for. Regarding time series of completely diffuse signals with some kind of trend the test would certainly fail to find a significant regime shift. However, in the rather shallow GWT time series evaluated here the breakpoints observed in the long term mean GWT are still statistically significant (p-values < 0.01), despite their, to a certain degree, diffusive nature. Thus, we are certain that our data fulfil the mentioned requirements and that the results are statistically sound.

Comment #23: p. 3653 and 3654: See specific comments on modelling and its interpretation.

Reply #23: This section was extensively modified. Please see our replies to the specific comments #8-10.

Comment #24: p. 3655, lines 17-20: This statement should be removed/modified. Results of the present study do not support this.

Reply #24: This sentence was modified to remove any reference to GWT damping in comparison to SAT warming. This has been shown to be the case in other studies, but admittedly was not demonstrated in the present one.

Comment #25: Conclusions should be more geared to specific findings/result of the study and not repeat discussion items that are not true results / have not been explicitly demonstrated.

> **Reply #25:** We agree. We removed the sentences at the end of the conclusions that discussed GDE (this information will be retained in the discussion) and included a sentence that states that groundwater temperature evolution in response to climate change can be tracked by analyzing long term records of the temperature of pumped groundwater (p. 18, lines 7-10).

Comment #26: Referencing is generally fairly complete. A modelling study dealing with the same time period and various factors influencing subsurface signals, including groundwater flow, heterogeneity and surface influences that may be of some use: Earth and Planetary Science Letters, 270, 86-94**.**

> **Reply #26:** Kooi (2008) will be added as a citation in the introduction (p. 3, lines 30-32) and in the reference list (p. 21, lines 10-12).

Reply to review by Y. Fan:

Comment: I find it interesting that the water table is deeper at Dansweiler and Sinthern (~17m), at a lower elevation of ~61m, than the Hardtwald wells (water table ~6m), although at a much higher elevation of ~121m. The two sites are not too far apart, and unless the underlying geology and hydrogeology (sources/sinks) make them entirely separate flow systems, one would expect that the water table is shallower in the lower part of the landscape. Perhaps the authors could discuss the hydrogeological system in more details and provide insights on this.

> **Reply**: The two study sites close to Karlsruhe and Cologne are approx. 240 km apart from each other and belong to different aquifer systems, which cannot be seen from Figure 1. The description of the hydrogeological settings was extended accordingly to clarify this issue (p. 5, lines 3-6).

Comment: Equation-1. The bulk thermal diffusivity is missing.

> **Reply**: Correct. In the submitted version of the manuscript the bulk thermal diffusivity was indeed missing. This error was corrected during type-setting and in the online HESSD version of the manuscript Equation 1 is displayed correctly including the thermal diffusivity (p. 7, line 23).

Comment: Equation-2. Perhaps the authors should mention that the Darcy velocity q in the convective term is related to water table recharge rate.

> Reply: We agree. A statement that the *q* term (Darcy velocity) is taken as the recharge rate was included in the manuscript below Equation 2 (p. 7, lines 29-30).

Comment: Page-11, line 33. Remove one of the “in”.

> **Reply**: The second ‘in’ was removed (p. 13, line 3).

Comment: Page-12, line 21-24. I tend to think that GWT change should be more as a trend than as steps, because of its delay and dampening of atmospheric signals, which smooth out the sharp rises (and falls) of the surface forcing. So a more fundamental explanation is perhaps not the short regime, but the nature of groundwater response. This bears out in the p values anyway, as the last sentence of the paragraph suggested.

> **Reply**: We agree that the general diffusive nature of the subsurface thermal signal in comparison to the air or surface thermal signals should be discussed more prominently in this section. We rephrased the paragraph accordingly, not only noting possible

methodological reasons for this significant trend here (p. 13, lines 22-23), but also mentioning the generally more gradual behavior of the GWT regimes (p. 13, lines 16-20, lines 25-26).

Comment: Page-14, line 18 and onward. Maybe replace the word “annual” to “inter-annual” to avoid confusion with “annual cycle” which refers to seasonal cycle within a year, not between years as is the case here? There are more of this word later, e.g., page-16, line-11.

Reply: We agree. The word ‘annual’ may indeed be misleading in this context, as we refer to the variations in the annual values. Consequently, the respective words were changed into ‘inter-annual’ (p. 15, lines 13-14; p. 18, line 1).

Observed groundwater temperature response to recent climate change

K. Menberg[1,2], P. Blum[1], B.L. Kurylyk[3] and P. Bayer[2]

[1] Karlsruhe Institute of Technology (KIT), Institute for Applied Geosciences (AGW), Karlsruhe, Germany.

[2] ETH Zurich, Department of Earth Sciences, Zurich, Switzerland.

[3] University of New Brunswick, Department of Civil Engineering and Canadian Rivers Institute, Fredericton, NB, Canada.

Correspondence to: P. Blum (philipp.blum@kit.edu)

Keywords: groundwater temperature, climatic change, regime shift, aquifer thermal regime, groundwater-dependent ecosystems

Abstract

Climate change is known to have a considerable influence on many components of the hydrological cycle. Yet, the implications for groundwater temperature, as an important driver for groundwater quality, thermal use and storage, are not yet comprehensively understood. Furthermore, few studies have examined the implications of climate change-induced groundwater temperature rise for groundwater-dependent ecosystems. Here, we examine the coupling of atmospheric and groundwater warming by employing stochastic and deterministic models. Firstly, several decades of temperature time-series are statistically analyzed with regard to climate regime shifts (CRS) in the long-term mean. The observed increases in shallow groundwater temperatures can be associated with preceding positive shifts in regional surface air temperatures, which are in turn linked to global air temperature changes. The temperature data are also analyzed with an analytical solution to the conduction-advection heat transfer equation to investigate how subsurface heat transfer processes control the propagation of the surface temperature signals into the subsurface. In three of the four monitoring wells, the predicted groundwater temperature increases driven by the regime shifts at the surface boundary condition generally concur with the observed groundwater

temperature trends. Due to complex interactions at the ground surface and the heat capacity of the unsaturated zone, the thermal signals from distinct changes in air temperature are damped and delayed in the subsurface, causing a more gradual increase in groundwater temperatures. These signals can have a significant impact on large-scale groundwater temperatures in shallow and economically important aquifers. These findings demonstrate that shallow groundwater temperatures have responded rapidly to recent climate change and thus provide insight into the vulnerability of aquifers and groundwater-dependent ecosystems to future climate change.

1 Introduction

Atmospheric climate change is expected to have a significant influence on subsurface hydrological and thermal processes (e.g. Bates et al., 2008; Green et al., 2011; Gunawardhana and Kazama, 2012). While the consequences for groundwater recharge and water availability were scrutinized by many studies (e.g. Maxwell and Kollet, 2008; Ferguson and Maxwell, 2010; Stoll et al., 2011; Taylor et al., 2013; Kurylyk and MacQuarrie, 2013), the implications of changing climate conditions for the long-term evolution of shallow groundwater temperatures are not comprehensively understood (Kløve et al., 2013). Groundwater temperature (GWT) is known to be an important driver for water quality (e.g. Green et al., 2011; Sharma et al., 2012; Hähnlein et al., 2013) and therefore, it is a crucial parameter for groundwater resource quality management (Figura et al., 2011).

Furthermore, increasing groundwater temperatures can have a significant influence on groundwater and river ecology (e.g. Kløve et al., 2013). Numerous studies on the impact of recent or projected climate change on the thermal regimes of surface water bodies and the associated impact for coldwater fish habitats have already been conducted (e.g. Kaushal et al., 2010; van Vliet et al., 2011, 2013; Wenger et al., 2011; Isaak et al., 2012; Wu et al., 2012; Jones et al., 2014), but the thermal sensitivity of shallow aquifers to climate change is a relatively unstudied phenomenon (e.g. Brielmann et al., 2009, 2011; Taylor and Stefan, 2009; Kurylyk et al., 2013, 2014a). The thermal response of GWT to climate change is of particular interest to river temperature analysts, as the thermal regimes of baseflow-dominated streams or rivers and hydraulically connected aquifers are inextricable linked (Hayashi and Rosenberry, 2002; Tague et al., 2007; Risley et al., 2010). Furthermore, groundwater sourced coldwater plumes within river mainstreams are known to provide thermal refuge for threatened coldwater fish (e.g. Ebersole et al., 2001; Breau et al., 2007), and questions have arisen regarding the sustainability of these groundwater-dependent ecosystems (GDEs) in a warming climate (Deitchman and Loheide, 2012). The current lack of knowledge regarding the thermal vulnerability of GWT to climate change and the associated impacts to GDEs has been highlighted as a research gap in several recent studies (e.g. Bertrand et al., 2012; Mayer, 2012; Kanno et al., 2014).

Thermal signals arising from changes in ground surface temperatures (GST) propagate downward into the subsurface, causing GWT to deviate from the undisturbed geothermal gradient. Heat transport theory has been applied for inverse modeling of temperature-depth profiles to infer paleoclimates based on measured deviations from the geothermal gradient (e.g. Mareschal and Beltrami, 1992; Pollack et al., 1998; Beltrami et al., 2006; Bodri and Cermak, 2007) and for forward modeling the impact of projected climate change on measured temperature-depth profiles (e.g. Gunawardhana and Kazama, 2011; Kurylyk and MacQuarrie, 2014). Such studies are often based on the assumption that long term trends in GST will track long term trends in surface air temperature (SAT), although this has been a matter of considerable debate (e.g. Mann and Schmidt, 2003; Chapman et al., 2004; Schmidt and Mann, 2004). For example, decreases in the duration of thickness of the insulating winter snowpack due to rising SAT can paradoxically lead to decreased winter GST (Smerdon et al., 2004; Zhang et al., 2005; Mellander et al., 2007; Mann et al., 2009; Kurylyk et al., 2013), which lead to a decoupling of mean annual SAT and GST trends.

Heat advection due to groundwater flow may also perturb subsurface temperature-depth profiles, and it can be difficult to determine if deviations from a linear geothermal gradient have arisen from past climate change or from groundwater flow (Reiter, 2005; Ferguson and Woodbury, 2005; Ferguson et al., 2006). Thus, several analytical solutions have been proposed that account for subsurface thermal perturbations arising from a combination of climate change and vertical groundwater flow (e.g. Taniguchi et al., 1999a, b; Kurylyk and MacQuarrie, 2014). The solutions vary depending on the nature of the surface boundary conditions employed (e.g. linear, exponential, or step trends in temperature), which can be used to match measured or predicted GST trends for a region. These solutions do not account for horizontal groundwater flow, which can also perturb subsurface thermal regimes in certain environments (Ferguson and Bense, 2011; Saar, 2011). Numerical solution techniques can also be applied to account for inhomogeneous subsurface thermal properties, complex surface temperature evolution, and groundwater flow (e.g., Kooi, 2008).

Figura et al. (2011) show that temperature variations in Swiss aquifers that are recharged by river water through bank infiltration can be related to changes in climate oscillations systems by applying a statistical regime shift analysis. Characterizing changes in time-series of various climatic, physical and biological parameters with the concept of abrupt regime shifts has been the focus of numerous studies in the last two decades (e.g. Hare and Mantua, 2000; Overland et al., 2008). In this context, a regime is often defined as a period with quasi-stable behavior or with a quantifiable quasi-equilibrium state (deYoung et al., 2004), and accordingly a rapid transition between states with differing average characteristics over multi-annual to multi-decadal periods is referred to as a regime shift (Bakun, 2004).

In this study, we demonstrate the direct influence of atmospheric temperature development on shallow GWT at two sites in Germany by analyzing time-series of SAT and GWT with regard to abrupt changes in the long-term annual mean. Compared to previous studies, which used borehole temperature profiles for the analysis of temperature coupling between the atmosphere and the subsurface, the measured time series of annual GWT of the last decades in this study allow for an evaluation of this coupling on a shorter time scale with a higher temporal resolution. Furthermore, we compare different spatially averaged temperature time-series from individual weather stations to global mean air temperature change bringing our observations in the context of global climate change. The magnitudes of the regime shifts and the time lags between the shifts in the chosen time-series are evaluated under consideration of the different thermal processes in the subsurface and the site-specific hydrogeological settings. A standard analytical solution to the conduction-advection subsurface heat transfer equation is applied to investigate the physical thermal processes underlying the observed correlation between SAT regime shifts and GWT rise.

2 Data and methods

2.1 Data and site description

For the analysis of shallow GWT, we use time-series from four observation wells in porous and unconfined aquifers in Germany (Table 1, Fig. 1a and b). Two of the wells are installed in the surrounding area of Cologne outside the small villages of Dansweiler and Sinthern in agricultural areas. The other two wells are located in a rather densely vegetated forest, called Hardtwald, close to the city of Karlsruhe and are therefore named Hardtwald 1 and 2. The proximate surroundings of all four wells were undisturbed over the last decades, so that variations in GWT due to land use changes are unlikely. The distances from the observation wells to the nearest streams are several kilometers (Table 1), thus the influence of river water on the groundwater temperature in the wells can be excluded. The two study areas close to Karlsruhe and Cologne are located approx. 240 km apart from each other and belong to different aquifer systems. Yet, the basic geological and hydrogeological settings of the two aquifers are rather similar (Table 1 and 2).

Table 2 lists some basic hydrogeological properties of the studied aquifers and the observation wells. The depth of water table differs considerably between the two well fields, and is approximately 17m for the Cologne aquifer and 7m near Karlsruhe. Variations in the depth of water table during the observation period are within ±1m for the Dansweiler and Sinthern wells and more pronounced in the Hardtwald wells with about ±3 m, which are likely caused by a pumping station nearby. However, no statistically significant trend was observed over the last decades in the water level of the observation wells. Both aquifers are recharged by infiltration of meteoric water through the unsaturated zone with estimated recharge rates of 221±45mmyr^{-1} for the Cologne aquifer and 228±45mmyr^{-1} for the aquifer near Karlsruhe (Table 2). A schematic cross-section of the two aquifers near Cologne (left) and Karlsruhe (right) in Fig. 1c and d shows the average depth of the water table below surface level and the depth of the underlying aquitard. Details on the wells' constructions are also depicted with the overall depth and the locations of the filter screens (black areas) that indicate the depth where the pumped water is captured. Furthermore, Fig. 1c and d shows the distance between the wells pairs as well as the distances to the weather stations from which the SAT time-series were obtained.

GWT in all observations wells was measured one to six times per year for a period of at least 32 years (1974–2006) during frequent water quality assessments by the local groundwater authorities. The measurement protocol, which is standardized by the environmental state agencies to assure data quality and comparability, has undergone no significant changes in the last decades. During the specified procedure, water is pumped from the wells until the water temperature and other on-site parameters are constant. The temperature measurements are thereby conducted with a probe directly at the outlet, to minimize influences on ambient air temperatures. An examination of the time series for seasonal effects revealed that they contain certain minor seasonal effects with variations, which indicates an impact of ambient air temperature on the GWT during the sampling. However, the natural temperature variations due to seasonal GST variations in depths of over 20 m (Table 2) are expected to be less than 0.1 K as can be demonstrated by Stallman's (1965) equation. In most years, at least two measurements per year were available, so that the arithmetic mean can be adopted as an annual mean value to minimize such effects. It should also be noted that the measurement accuracy is in the range of ±0.1 K. Also changes in the measurement procedure, such as variations in the pumping rate or in the placement of the pump within the well, as well as changes in the measurement equipment, can influence the measured GWT and should be considered for the evaluation and interpretation of the data.

Annual SAT data are available from weather stations operated by the German Weather Service (DWD) outside the cities of Cologne and Karlsruhe in agricultural surroundings (Fig. 1a and b). Though located several kilometers from the observation wells, the SAT from these stations is expected to yield a good approximation for the development of SAT at the well sites. Furthermore, for the evaluation of abrupt shifts in the time series of SAT and GWT, the absolute temperature is only of minor importance, while the main focus is on the timing of the shifts and the temperature differences. For the comparison with air temperatures on a larger scale, we use time-series of mean air temperature anomalies based on the reference period 1951–1980 from the NASA Goddard Institute for Space Studies (GISS) (e.g. Hansen et al., 2010). Of the spatially averaged temperature data sets available, we evaluate the annual global mean from land-surface air and sea-surface water temperature anomalies and the annual zonal mean for the Northern Hemisphere between 90° and 24°N based on land-surface air temperature anomalies.

2.2 Regime shift analysis

There are several possibilities to statistically evaluate temperature changes in time series with rather simple functional forms. Seidel and Lanzante (2004) compared different approaches (e.g. linear and flat steps models) and revealed that often time-series of atmospheric temperatures can be represented more appropriately by models using breakpoints than by models assuming monotonic functions. Hence, we here apply a sequential t-test analysis for regime shifts (STARS) to detect possible abrupt regime shifts (CRS) in the temperature time-series (Rodionov, 2004; Rodionov and Overland, 2005). The STARS method has been successfully used by recent studies to identify abrupt changes in the long-term mean of environmental time-series (Marty, 2008; North et al., 2013) and GWT time-series (Figura et al., 2011). STARS is a parametric test that can detect multiple regime shifts and needs no a priori assumption for the timing of possible shifts. Identification of a shift is based on the calculation of the Regime Shift Index (RSI), which represents the cumulative sum of the normalized deviations from the mean value of a regime and thus reflects the confidence of a regime shift (Rodionov, 2004). For the regime shift analysis, several test parameters need to be adjusted to account for specific characteristics, such as the length of the tested time-series. The target significance level in our analysis is set to 0.15, which corresponds to the p-level of false positives. The actual p value of an identified shift between subsequent regimes is calculated separately with a Student's t test. The cut-off length of the test corresponds to a low-pass filter, so that regimes with a shorter length are disregarded in the analysis (Rodionov and Overland, 2005). Here, we set the cut-off length to 10 years as atmospheric oscillations often occur at decadal intervals (Overland et al., 2008). Furthermore, the Huber weight parameter (set to 1 in our study) included in the STARS procedure improves the treatment of outliers by weighting them proportionally to their deviation from the mean value (Overland et al., 2008). As pointed out by Seidel and Lanzante (2004) atmospheric data tend to be highly temporally auto-correlated, so that especially in short time-series, spurious regime shifts may be detected due to serial correlation (Rudnick and Davis, 2003). Therefore, we apply a pre-whitening procedure that removes the red noise component from the temperature time series prior to testing for a regime shift (Rodionov, 2006). To investigate the potential stationarity within detected regimes, the non-parametric Mann-Kendall test for the absence of trend is also applied to the temperature data (von Storch, 1995).

2.3 Analytical solutions

The governing equation for transient subsurface heat transport is the one dimensional conduction equation for homogeneous media, which equates the divergence of the conductive flux with the rate of the change of thermal energy in the medium (Carslaw and Jaeger, 1959; Domenico and Schwartz, 1990):

$$\kappa \frac{\partial^2 T}{\partial z^2} = \frac{\partial T}{\partial t} \quad (1)$$

where κ is the bulk thermal diffusivity of the subsurface ($m^2\ s^{-1}$), T is temperature (°C), z is depth (m), and t is time (s). The governing heat transport equation becomes slightly more complex when advective heat transport (or 'forced convection') due to groundwater flow is considered:

$$\kappa \frac{\partial^2 T}{\partial z^2} - U \frac{\partial T}{\partial z} = \frac{\partial T}{\partial t} \quad (2)$$

where U ($m\ s^{-1}$) is the thermal plume velocity under pure advection and a function of the Darcy velocity q (downwards or recharge is positive, $m\ s^{-1}$), the bulk volumetric heat capacity of the soil-water matrix C (J m^{-3} °C^{-1}), and the volumetric heat capacity of water C_w (J m^{-3} °C^{-1}):

$$U = q\frac{C_w}{C} \quad (3)$$

The governing conduction-advection equation (2) employs several limiting assumptions, including spatiotemporally constant groundwater velocity over the entire domain (including depths below the well screen), one dimensional heat transport, homogenous thermal properties, constant pore water phase, and isothermal conditions between the soil grains and pore water. Here we employ a distinct analytical solution to Eq. (2) to simulate the influence of a climate regime shift on GWT. We assume thermally uniform initial conditions and boundary conditions that are subject to a series of n step increases in GST:

Initial conditions: $T(z,t=0)=T_0$ (4)

Boundary condition: $T(z=0,t)=T_0+\sum_{i=1}^{n}\Delta GST_i \times H(t-t_i)$ (5)

where T_0 is the initial uniform temperature (°C) prior to the beginning of the regime shift, ΔGST_i is the step increase in GST for regime shift i (°C), H is the Heaviside step function, and t_i is the time (s) of the beginning of regime shift i. In this formulation, ΔGST_i refers to a step change in GST in comparison to the GST conditions immediately preceding that change (not necessarily in comparison to initial GST, T_o). We ignore short term (e.g. annual) variations in SAT and GST and rather drive the subsurface heat transport models with temperatures averaged for a given climate regime and then instantaneously increased at the beginning of the next climate regime. The thermally uniform initial conditions is a reasonable assumption given that we begin by considering mean annual GWT at or near the water table following a relatively stable climate regime (i.e. prior to 1988, Fig. 2). Moreover, for the wells observed, the vadose zones and near-surface aquifers are too shallow to realize the influence of any geothermal gradient. The isothermal condition assumption previously noted extends to the surface boundary, which implies that the groundwater recharge entering the semi-infinite domain at the ground surface has a temperature equal to the mean annual surface temperature for that climate regime.

The transient conduction-advection heat transport model (TCA model) employed in this study is an analytical solution to the transient conduction-advection Eq. (2) subject to the initial and boundary conditions given in Eqs. (4) and (5). This solution was originally developed by

Carslaw and Jaeger (1959) and subsequently employed by Taniguchi et al. (1999b) to study subsurface temperature evolution due to land cover changes in regions of significant groundwater flow. Because we assume initially thermally uniform conditions in the unsaturated zone and shallow groundwater, the resultant solution is simpler than in the original derivations. Unlike the original derivation, it is also presented here with superposition principles applied to allow for a series of regime shifts rather than one event. This superposition approach is valid given the linearity of the governing partial. This superposition approach is valid given the linearity of the governing partial differential equation and the boundary and initial conditions (Farlow, 1982):

$$T(z,t) = T_0 + \sum_{i=1}^{n} \frac{\Delta GST_i}{2} \left\{ erfc\left(\frac{z - U(t-t_i)}{2\sqrt{\kappa(t-t_i)}} \right) + \exp\left(\frac{Uz}{\kappa} \right) erfc\left(\frac{z + U(t-t_i)}{2\sqrt{\kappa(t-t_i)}} \right) \right\} \times H(t-t_i) \quad (6)$$

where *T(z,t)* is the spatiotemporally varying subsurface temperature (GWT, °C), κ is the bulk thermal diffusivity of the subsurface (m^2 s^{-1}), and erfc is the complementary error function. The Heaviside function indicates that the subsurface thermal influence of each regime shift *i* in the boundary condition is not realized until the time *t* exceeds t_i. Comparisons between the model results and measured GWT indicate whether these simple analytical solutions are applicable for modeling the influence of observed and projected climate regime shifts in the wells considered in this study.

It should be noted that it is the GST rather than the SAT that drives subsurface thermal regimes and thus forms the boundary condition in Eq. (5). However, complete GST time series were not available for the locations considered in this study. Thus, in the present study, the magnitude and timing of the regime shifts in GST are obtained from the local SAT data as follows. In all cases, the timing of the GST regime shifts is assumed to correspond to the timing of the local SAT regime shifts for that location obtained from the statistical analysis. This approach is reasonable given the efficient heat transfer that occurs between the lower atmosphere and the ground surface (e.g. Bonan, 2008). The magnitude of the GST regime shift was set to be equal to the magnitude of the SAT. Measured SAT and GST data (not shown) indicate that this approach is valid as the measured magnitude of the climate regime shift in 1988 was 1.1 C in both the SAT and GST data near Cologne. No GST data were available for the sites near Karlsruhe (Hardtwald sites, Figure 1 and Table 2), but it is reasonable to assume that the magnitude of the GST changes track the magnitude of the SAT changes like in the case of the sites near Cologne.

Table 3 presents the assumed subsurface thermal properties for each well for both the saturated and unsaturated zones. A potential range in these values was estimated from literature values taking into account variations in lithology obtained from drilling logs as wells as the variability of water content in the unsaturated zone ranging from dry to saturated conditions (VDI, 2010; Menberg et al., 2013). Because we consider temperature rise within the aquifer, the effective thermal diffusivities utilized in the analytical solution for each of the four locations were obtained from a weighted arithmetic average (weighted by zone thickness) of the saturated and unsaturated zone thermal diffusivities. For example, the unsaturated zone thickness was taken as the depth to the water table, and the saturated zone thickness was taken as the distance from the water table to a point along the well screen. Different points in the well screen were considered, as described in the results, because the vertical well capture zone flow dynamics may be complex depending on the nature of the pumping and heterogeneities in near-well hydraulic properties. This is particularly important for the Hardtwald wells, which have longer well screens than in the case of the Dansweiler or Sinthern wells (Fig. 1).

Regional recharge rates were extracted from Table 2 with a potential range to reflect the variability of recharge in this region over the last decades (Erftverband, 1995; W. Deinlein, personal communication, 2013). Similar thermal properties and recharge values are assumed for Hardtwald 1 and Hardtwald 2 based on their similar land cover and subsurface properties and the geographical proximity (about 200 m) between the wells.

3 Results and Discussion

3.1 Statistical analysis

3.1.1 Regime shifts in air and groundwater temperatures

At least two climate regime shifts (CRS) could be detected in the later decades of all analyzed time-series (Fig. 2). The time-series of global mean temperature change and zonal mean temperature change in 90-24° N show significant (STARS, $p < 0.005$) positive shifts in 1977, 1987, 1997 and 1977, 1988 and 1998, respectively (Table 4). The observation of shifts in air temperature change in these years is in good agreement with the observation of decadal shift in atmospheric oscillation indices in the late 1970s, late 1980s and late 1990s (Overland et al., 2008). Only the CRS in the late 1980s and late 1990s can be found from examining the time-series of local SAT data from Cologne and Karlsruhe. However, this is not surprising as previous studies observed that the CRS in the late 1970s was most prominent in the North Pacific region (Hare and Mantua, 2000; Overland et al., 2008), and less accentuated in Europe. The same applies to the CRS in the late 1990s (Overland et al., 2008; Swanson and Tsonis, 2009), which is reflected by the differing RSI values in Fig. 2. While the high RSI for the CRS in 1997 in the global mean temperature change indicates a significant shift, the RSIs for the late 1990s CRS in the German SAT time-series are much lower than the RSIs in the late 1980s. Figura et al. (2011) correlated the abrupt increase in SAT in Switzerland with a change in the Artic Oscillation (AO) that has a strong influence on air temperatures in Europe. However, no such change in the AO Index was found in the late 1990s, suggesting that the CRS in the German SAT is also coupled to the general air temperature increase in the Northern Hemisphere.

Two regime shifts were detected in the GWT time-series for the four wells near Cologne and Karlsruhe. These shifts correspond to the CRS in the atmosphere with a certain time lag (Fig. 2, Table 4). The regime shifts in GWT time-series are all statistically significant ($p < 0.01$), except for the second regime shift in the late 1990s in Dansweiler. Two prominent outliers in the third regime of the time-series influence the statistical significance for this shift, while the RSI value is calculated under consideration of the outliers according to the Huber weight parameter. Furthermore, the RSI values in Fig. 2 for the second shifts in Dansweiler and Sinthern are not the final values, as the 10-year cut-off length of the STARS test in the last regime has not yet been reached. In general, the time-series of GWT show a more gradual increase than the SAT time-series. In particular, the GWT in the Sinthern well appears to exhibit a linear trend rather than a step increase, which is subsequently discussed. The GWT time-series partly exhibit considerable inter-annual variability, which appears to be more significant in the Hardtwald wells than in Dansweiler and Sinthern. Potential reasons for these rather large fluctuations in annual GWT are related to the uncertainties associated with the measurements as mentioned in the method section. Other possible factors that influence the inter-annual variability could be the pumping station close to the wells in the Hardtwald, where groundwater is extracted at irregular intervals, and impacts by undetected land use changes in the close surroundings.

3.1.2 Statistical analysis of time lags and magnitude of temperature change

The time lags between the regime shifts in SAT and GWT are listed in Table 4. The regime shifts in global mean temperature change and the zonal mean in 90-24° N occur simultaneously, except for the regime shift in the late 1980s that has a time lag of one year. However, as annual mean values are used for the analysis, the accuracy of the shift detection is limited to ±2 year, so that the shifts occur within the uncertainty range. The same applies to the first regime shifts in the local SAT time-series in Cologne and Karlsruhe. A possible explanation for this variation in the time lags would be that the late 1980s regime shift was very prominent in the Arctic Oscillation that directly influences the European climate (Figura et al., 2011). The late 1990s regime shift however, was more distinct in the North Pacific region (Overland et al., 2008), thus probably causing the delayed shift in the SAT in Germany. Furthermore, changes in local SAT are also expected to be temporally and spatially highly heterogeneous due to the variability of local climate and the complexity of atmospheric circulation systems (Hansen et al., 2010). The observed CRS in shallow GWT lag behind the abrupt increase in local SAT by 1–4 years (Table 4). In Karlsruhe the time lag is generally small with one year for all shift events, while the time lags in Cologne vary between 2–4 years. This difference in the time lags reflects the specific hydrogeological site conditions with the unsaturated zone in Cologne (17m) being significantly thicker than in Karlsruhe (7m, Table 2). The thermal properties in the unsaturated zone differ significantly from those in the saturated zone (Table 3). Thus the propagation of the thermal signal in Cologne is retarded due to the lower thermal diffusivity than in Karlsruhe.

The magnitudes of the temperature increase between two subsequent regimes in the zonal mean SAT change are considerably higher than in the global mean SAT change (Fig. 3), because the global temperature data set contains ocean temperature measurements, and ocean temperatures are known to respond more slowly to climatic forcing due to the ocean's large thermal inertia (Hansen et al., 2010). The above mentioned temporal and spatial heterogeneity of the CRS accounts also for the higher increase in SAT in the German time series, which is above the average of the zonal mean in 90-24° N. The significant abrupt increase in the long-term mean of SAT with the late 1980s CRS of close to 1°C was likewise observed in Swiss SAT by Figura et al. (2011).

The magnitudes of the increases in the long-term means of GWT are lower and damped by up to 70% compared to the shift magnitude in SAT (Fig. 3). This damping arises from the fact that, due to the thermal inertia of the subsurface, the GWT has not yet fully equilibrated with the GST at the time when the regime shift is observed in the GWT. The magnitudes of the regime shifts in Fig. 3 also reveal that the damping in the time-series from the Hardtwald wells is more pronounced than the damping in Dansweiler and Sinthern. This likely occurs due to depth of groundwater extracted for temperature measurements. For example, the depth to the midpoint of the well screen is higher for the Hardtwald wells than it is for the Dansweiler and Sinthern wells (Table 2). This will be investigated in more detail with the TCA model.

3.1.3 Stationarity within the regimes

In order to investigate the stationarity within the identified regimes the Mann-Kendall test for the absence of trend was performed for the individual regimes. The resulting p-values are listed in Table 5, in which high p-values close to 1 indicate stationary conditions. No significant trends could be found within the individual regimes of the examined SAT time-series, suggesting that the temperature increase in the last decades can be attributed completely to the detected CRS.

In the GWT time series, the p-values of the Mann-Kendall test are generally lower (median of 0.20, Table 5) than the p-values of the SAT time series (median of 0.53), indicating that the SAT time-series are more stationary than GWT time series. This more gradual increase in GWT reflects the effects of subsurface heat transport dynamics, which convert the sharp surface temperature signal to a more diffuse subsurface temperature signal. A significant trend ($p < 0.05$) with a slope of 0.13°C was detected in the third regime (2001–2006) in the Sinthern well. However, it has to be noted, that this regime is quite short, and thus the trend analysis may be biased by the last two rather high temperature values in 2005 and 2006. In the regimes before 1991, the p-values of the time-series in Dansweiler and Sinthern are 0.05 and 0.06, respectively, and thus close to the critical p-value of 0.05 indicating a more gradual increase rather than abrupt changes. For the wells near Karlsruhe no significant trends were found in GWT within the regimes, which indicates that the temperature increase in the time-series can be linked to the regime shifts.

To compare the performance of the regime shift analysis to an approach with linear temperature increase, the RMSE values for the statistical step function model and a linear model were calculated for each time series (not shown). This analysis revealed that the RMSE of the step function fit for all GWT and SAT time series is slightly lower than the RMSE of the linear fit, indicating that the step function model performs slightly better. Thus, it can be stated, that, with the exception of the potentially biased last regime in Sinthern, all regimes in the time series of GWT are statistically stationary, which corroborates the feasibility of the application of regime shift analyses on GWT time-series in addition to the low p-values of STARS (Table 4).

3.2 Analytical model

Predicted GWT were obtained from the analytical solution in Eq. (6) (TCA model) with the thermal properties and recharge rates given in Table 3 and the magnitude and timing of the regime shifts given in Table 4. Due to the availability of GWT data in each well, model runs were started in 1970. Figure 4 shows the measured GWT, assigned GST boundary condition, and predicted GWT for each of the four wells. The range of predicted GWT (shaded area, Fig. 4) is derived from the range of thermal properties, well screen depths, and recharge values utilized as input parameters to the model (Table 3). In particular, the upper boundaries of the temperature envelopes in Fig. 4 were obtained with Eq. (6) using depths to the tops of each well screen (Table 2), maximum thermal diffusivities (Table 3), and minimum heat capacities (Table 3, see Eq. 3). The lower boundaries of the temperature envelopes were obtained using depths to the bottom of the well screens, minimum thermal diffusivities, and maximum heat capacities. Finally, the best estimates (red series, Fig. 4) for the predicted GWT data for each well were obtained using depths equal to the midpoints of the well screens and mean thermal diffusivity and heat capacity. In all cases, the thermal properties were taken as the weighted arithmetic average of the unsaturated and saturated zone thermal properties as described in section 3.2.

Note that the GST data simulated for the Hardtwald wells are characterized by a wider range in the predicted temperature envelopes. This range is primarily due to the longer well screens in the case of the Hartwald wells than for the Sinthern and Dansweiler wells (Table 2). Hereafter, when we refer to the TCA model results we ignore the range in the modeling results and only allude to the specific results obtained using the mean recharge values and mean thermal properties given in Table 3 (i.e. red series, Fig. 4).

The TCA model predicted trends in GWT generally concur with the long term trends exhibited in the measured data for Dansweiler, Hartwald 1, and Hartwald 2; however, the TCA model under-predicts the rise in the Sinthern GWT data. These differences suggest that, although they were assumed to be equal, the magnitude of the GST regime shifts in Sinthern may have been greater than those in Dansweiler, or that due to subsurface heterogeneity, the pumped water may be predominantly sourced from above the Sinthern well screen. Modeling results (not shown) indicate that for the mean thermal properties and recharge values, the *z* value used in Eq. (6) would have to be approximately 9 m for the predicted and observed GWT trends to generally concur. Furthermore, the recharge rates in this well may have been greater than the obtained regional recharge rates for this area. Higher recharge would lead to higher heat advection, which would reduce the lag between a GST signal and its realization in the subsurface (see range in predicted Sinthern GWT, Fig. 4). Similarly, higher thermal diffusivity would generally lead to higher GWT in Sinthern, as the Sinthern GWT is still adjusting to the GST regime shifts in the data shown in Fig. 4. Finally, the last few years of measured GWT data are not available for the Sinthern well. GWT data in the nearby Dansweiler well decreased during this period, thus the visual fit between the measured and predicted Sinthern GWT would likely improve if these data were available.

Our approach does not reproduce inter-annual variability in GWT due to the nature of the GST boundary condition, which is constant for a given climate regime (Fig. 4). Inter-annual variability in GWT could theoretically be reproduced by considering a series of "GST regimes" that only last one year; however, the objective of the present study was to examine the subsurface thermal influence of climate regime shifts not inter-annual SAT or GST variability. Finally, it is interesting to note that the abrupt regime shifts applied in the simplified boundary condition manifest themselves as gradual changes in the predicted GWT evolution in the deeper wells due to the influence of the heat capacity and thermal inertia of the subsurface. These findings demonstrate that observed gradual increases in shallow GWT are not necessarily suggestive of gradual trends in GST. The effect of the abrupt GST regime shifts are discernible in the upper edge of the temperature envelopes in Hardtwald 1 and 2 (i.e., the GWT signal is diffused but the impact of the piecewise boundary condition is still discernible). This is due to the fact that these particular results were obtained for depths to the top of the well screens of only 10 m (Table 2).

With the exception of the anomalous Sinthern data, the general agreement between the predicted and observed trends in GWT data (Fig. 4) indicates that TCA model can produce first-order approximations of the thermal sensitivity of these shallow aquifers to past or future climate regime shifts by conforming the boundary condition to climate model projections. The boundary condition form employed in this study could be matched to future climate projections by considering a series of short GST regimes, or alternatively, a boundary condition that considers a gradual rise in GST could be employed (e.g., Kurylyk and MacQuarrie, 2014). The form of the analytical solution indicates that if a new long term stable climate is achieved, the GWT will eventually rise an equivalent magnitude to the changes in

GST, which are often in turn assumed to follow changes in SAT. In the absence of snowpack evolution or land cover changes, any perceived damping in GWT changes in comparison to SAT changes based on statistical analyses likely results from the lagged subsurface thermal response to the boundary condition.

There are limitations associated with employing analytical solutions to simple one-dimensional heat transport equations. Several assumptions associated with the conduction-advection equation have been previously noted. For example, the governing equation, and hence the analytical solution, assume that the water phase is constant. This assumption is justified in the present study considering that no permafrost thaw (which retards soil warming, Kurylyk et al., 2014b) is occurring. Also, the solution assumes homogeneous thermal properties; however, we considered heat transport in both the saturated and unsaturated zones. The thermal diffusivity of the unsaturated zone for the wells considered in the present study were up to 30% lower than the saturated zone thermal diffusivities (Table 3). We considered both zones by employing a weighted arithmetic average (based on zone depths) for the effective thermal diffusivity. Also, recharging water may be at a temperature different than the mean annual surface temperature, particularly if the recharge mechanism is snowmelt. However, snowmelt induced recharge is minimal at the observation areas in this study. In general, due to these limitations, the results presented in Figure 4 should be considered first order approximations of the sensitivity of these shallow aquifer thermal regimes to climate regime shifts.

3.2.1 Implications for future river temperatures and groundwater-dependent ecosystems

Although the wells analysed in this study were not located nearby streams, the timing and magnitude of the measured GWT rise can provide insight into the potential warming of alluvial aquifers feeding ecologically important rivers. Gaining rivers and streams can be strongly influenced by the thermal regimes of surrounding aquifers (e.g. Tague et al., 2007; Kelleher et al., 2012), and this is often particularly true during the dry, warm season when baseflow can provide the majority of the river or stream discharge. Thus, deterministic models of future base-flow dominated rivers temperature should explicitly account for the future thermal regimes of aquifers. Various studies have demonstrated that the thermal regimes of rivers respond to a warming climate, and these studies have generally tacitly ignored GWT rise due to climate change. The results of this study however contradict this assumption by indicating that shallow GWT will respond to SAT warming and that the lag time between

SAT warming and the associated increase in shallow GWT can be rather short (< 5 years). Similar results were obtained by Kurylyk et al. (2014a) who employed a numerical model of groundwater flow and energy transport driven by downscaled climate scenarios to demonstrate a potential damping and short lagging of future groundwater discharge temperature rise in response to air temperature changes.

Given the expected warming of rivers across the globe (van Vliet et al., 2011, 2013), researchers have rightfully proposed that coldwater fish will begin to increasingly rely on the occurrence and distribution of suitable coldwater refugia (e.g. Brewer, 2013). Our results suggest that GDEs and groundwater-sourced coldwater refugia will also warm in response to climate change. The magnitude and timing of the GWT warming however will depend on several factors, including the timing and magnitude of the SAT warming, changes in precipitation (and thus recharge and advection), the depth of the groundwater table, and the presence or absence of seasonal snowpack.

4 Conclusions

By applying a sequential t-test analysis for regime shifts (STARS) to time-series of air and groundwater temperatures, we empirically demonstrated that groundwater temperatures in shallow aquifers show temperature changes that correspond to positive shifts in local SAT in Germany, which in turn can be traced back to increasing global SAT. This observed direct coupling of atmospheric and groundwater temperature development through the unsaturated zone implies that climate warming does not only affect aquifers recharged by river-bank infiltration (Figura et al., 2011), but also a large number of shallow aquifers on a wide spatial scale. The regime shifts in GWT occur with a certain time lag to the CRS depending mainly on the thermal properties and thickness of the unsaturated zone. The magnitude of these regime shifts in GWT compared to the shifts in SAT is damped by the thermal propagation of the temperature signal into the subsurface, leading to a more gradual increase in GWT. This damping perceived in the statistical analyses is predominantly an artifact of the lagged subsurface thermal response. However, despite the extenuation of the temperature signal in the subsurface and the mixing of shallow groundwater during pumping, significant temperature shifts were found in the extracted groundwater.

Process-oriented modeling was also performed with an analytical solution to the conduction-advection equation. In three of the four observation wells, the simulated decadal GWT trends generally concurred with the measured decadal GWT trends, although inter-annual variability was not reproduced due to the simplistic nature of the boundary condition. This agreement is indicative that the solution to the conduction-advection equation can also be applied to obtain first-order estimates of the influence of future climate change on subsurface thermal regimes.

Our results indicate that increasing SATs are prone to have a substantial and swift impact, not only on soil temperatures, but also on large-scale, shallow groundwater temperatures in productive and economically important aquifers. Furthermore, this study has demonstrated that long-term series of pumped groundwater temperature can be analyzed using stochastic approaches to examine the relationship between local and global climate change and local groundwater temperature evolution.

Acknowledgements

We would like to thank Ying Fan and Henk Kooi for their constructive reviews. The financial support for K. Menberg from the Scholarship Program of the German Federal Environmental Foundation (DBU) is gratefully acknowledged. Furthermore, we would like to thank S. Simon (Erftverband), W. Feuerbach (Landesanstalt für Umwelt, Messungen und Naturschutz, LUBW), W. Deinlein (Stadtwerke Karlsruhe GmbH) and M. Kimmig for the support with long-term groundwater data. K. Sulston of the University of Prince Edward Island contributed valuable advice on the derivation of the analytical solution to the conduction-advection equation. B. Kurylyk received funding from Natural Sciences and Engineering Research Council of Canada postgraduate scholarships. We acknowledge support by the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of the Karlsruhe Institute of Technology.

References

Bakun, A.: Chapter 25 - regime shifts, in: The Sea, edited by: Robinson, A., and Brink, K., Harvard University Press, Cambridge, MA, 2004.

Balke, K.-D.: Geothermische und hydrogeologische Untersuchungen in der südlichen Niederrheinischen Bucht, Hannover, 1973.

Bates, B., Kundzewicz, Z. W., Wu, S., and Palutikof, J. P.: Climate Change and Water, Intergovernmental Panel on Climate Change Secretariat, Geneva, 210, 2008.

Beltrami, H., Gonzales-Rouco, J. F., and Stevens, M. B.: Subsurface temperatures during the last millenium: Model and observation, Geophys. Res. Lett., 33, L09705, doi:09710.01029/02006GL026050, 2006.

Bertrand, G., Goldscheider, N., Gobat, J.-M., and Hunkeler, D.: Review: From multi-scale conceptualization to a classification system for inland groundwater-dependent ecosystems, Hydrogeol. J., 20, 5-25, 2012.

Bodri, L., and Cermak, V.: Borehole climatology: A new method on how to construct climate, Elsevier, Amsterdam, 2007.

Bonan, G.: Ecological climatology, Cambridge University Press, United Kingdom, 2008.

Breau, C., Cunjak, R. A., and Bremset, G.: Age-specific aggregation of wild juvenile Atlantic salmon Salmo salar at cool water sources during high temperature events, J. Fish Biol., 71, 1179-1191, 10.1111/j.1095-8649.2007.01591.x, 2007.

Brewer, S. K.: Groundwater influences on the distribution and abundance of riverine smallmouth bass, micropterus dolomieu, in pasture landscapes of the midwestern USA, River Res. Appl., 29, 269-278, 2013.

Brielmann, H., Griebler, C., Schmidt, S. I., Michel, R., and Lueders, T.: Effects of thermal energy discharge on shallow groundwater ecosystems, FEMS Microbiology Ecology, 68, 242 - 254, 2009.

Brielmann, H., Lueders, T., Schreglmann, K., Ferraro, F., Avramov, M., Hammerl, V., Blum, P., Bayer, P., and Griebler, C.: Oberflächennahe Geothermie und ihre potenziellen Auswirkungen auf Grundwasserökosysteme, Grundwasser, 16, 77-91, 10.1007/s00767-011-0166-9, 2011.

Carslaw, H. S., and Jaeger, J. C.: Conduction of Heat in Solids (2nd ed.), New York, Oxford University Press, 510 pp., 1959.

Chapman, D. S., Bartlett, M. G., and Harris, R. N.: Comment on "Ground vs. surface air temperature trends: Implications for borehole surface temperature reconstructions" by M.E. Mann and G. Schmidt, Geophys. Res. Lett., 31, L07205 07201-07203, 2004.

Deinlein, W.: Stadtwerke Karlsruhe GmbH, personal communication, 2013.

Deitchman, R., and Loheide, S. P.: Sensitivity of thermal habitat of a trout stream to potential climate change, Wisconsin, United States, J. Am. Water Resour. Assoc., 48, 1091-1103, 10.1111/j.1752-1688.2012.00673.x, 2012.

deYoung, B., Harris, R., Alheit, J., Beaugrand, G., Mantua, N., and Shannon, L.: Detecting regime shifts in the ocean: Data considerations, Prog. Oceanogr., 60, 143-164, 10.1016/j.pocean.2004.02.017, 2004.

Domenico, P. A., and Schwartz, F. W.: Physical and chemical hydrogeology, second edition, 2nd ed., John Wiley & Sons, Inc., New York, NY, 842 pp., 1990.

Ebersole, J. L., Liss, W. J., and Frissell, C. A.: Relationship between stream temperature, thermal refugia and rainbow trout Oncorhynchus mykiss abundance in arid-land streams in the northwestern United States, Ecol. Freshwat. Fish, 10, 1-10, 10.1034/j.1600-0633.2001.100101.x, 2001.

Erftverband: Basisplan III zur Sicherstellung der Wasserversorgung im Bereich des Erftverbands, Erftverband, Bergheim, 1995.

Farlow, S. J.: Partial differential equations for scientists and engineers, Wiley, New York, 1982.

Ferguson, G., and Woodbury, A. D.: The effects of climatic variability on estimates of recharge from temperature profiles, Ground Water, 43, 837-842, 2005.

Ferguson, G., Beltrami, H., and Woodbury, A. D.: Perturbation of ground surface temperature reconstructions by groundwater flow?, Geophys. Res. Lett., 33, 1 - 5, 2006.

Ferguson, G., and Bense, V.: Uncertainty in 1D Heat-Flow Analysis to Estimate Groundwater Discharge to a Stream, Ground Water, 49, 336-347, 2011.

Ferguson, I. M., and Maxwell, R. M.: Role of groundwater in watershed response and land surface feedbacks under climate change, Water Resour. Res., 46, W00F02, 10.1029/2009wr008616, 2010.

Figura, S., Livingstone, D. M., Hoehn, E., and Kipfer, R.: Regime shift in groundwater temperature triggered by the Arctic Oscillation, Geophys. Res. Lett., L23401, 10.1029/2011GL049749, 2011.

Green, T. R., Taniguchi, M., Kooi, H., Gurdak, J. J., Allen, D. M., Hiscock, K. M., Treidel, H., and Aureli, A.: Beneath the surface of global change: Impacts of climate change on groundwater, J. Hydrol., 405, 532-560, 2011.

Gunawardhana, L. N., and Kazama, S.: Climate change impacts on groundwater temperature change in the Sendai plain, Japan, Hydrol. Processes, 25, 2665-2678, 2011.

Gunawardhana, L. N., and Kazama, S.: Statistical and numerical analyses of the influence of climate variability on aquifer water levels and groundwater temperatures: the impacts of climate change on aquifer thermal regimes, Global Planet. Change, 10.1016/j.gloplacha.2012.02.006, 2012.

Hähnlein, S., Bayer, P., Ferguson, G., and Blum, P.: Sustainability and policy for the thermal use of shallow geothermal energy, Energ. Policy, 59, 914-925, 10.1016/j.enpol.2013.04.040, 2013.

Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, Rev. Geophys., 48, RG4004, 8755-1209/10/2010GR000345, 2010.

Hare, S. R., and Mantua, N. J.: Empirical evidence for North Pacific regime shifts in 1977 and 1989, Prog. Oceanogr., 47, 103-145, 10.1016/S0079-6611(00)00033-1, 2000.

Hayashi, M., and Rosenberry, D. O.: Effects of ground water exchange on the hydrology and ecology of surface water, Ground Water, 40, 309-316, 10.1111/j.1745-6584.2002.tb02659.x, 2002.

HGK: Hydrogeologische Kartierung und Grundwasserbewirtschaftung im Raum Karlsruhe-Speyer, Umweltministerium Baden-Württemberg and Ministerium für Umwelt, Forsten und Versbraucherschutz Rheinland-Pfalz, Stuttgart, Mainz, 90, 2007.

Isaak, D. J., Wollrab, S., Horan, D., and Chandler, G.: Climate change effects on stream and river temperatures across the northwest US from 1980-2009 and implications for salmonid fishes, Climate Change, 113, 499-524, 10.1007/s10584-011-0326-z, 2012.

Jones, L. A., Muhlfeld, C. C., Marshall, L. A., McGlynn, B. L., and Kershner, J. L.: Estimating thermal regimes of bull trout and assessing the potential effects of climate warming on critical habitats, River Res. Applic., 30, 204-216, 10.1002/rra.2638, 2014.

Kanno, Y., Vokoun, J. C., and Letcher, B. H.: Paired stream-air temperature measurements reveal fine-scale thermal heterogeneity within headwater brook trout stream networks, River Res. Applic., 30, 745-755, 10.1002/rra.2677, 2014.

Kaushal, S. S., Likens, G. E., Jaworski, N. A., Pace, M. L., Sides, A. M., Seekell, D., Belt, K. T., Secor, D. H., and Wingate, R. L.: Rising stream and river temperatures in the United States, Frontiers in Ecology and the Environment, 8, 461-466, 10.1890/090037, 2010.

Kelleher, C., Wagener, T., Gooseff, M., McGlynn, B., McGuire, K., and Marshall, L.: Investigating controls on the thermal sensitivity of Pennsylvania streams, Hydrol. Processes, 26, 771-785, 2012.

Klostermann, J.: Das Quartär der Niederrheinischen Bucht—Ablagerungen der Letzten Eiszeit am Niederrhein, Geologisches Landesamt Nordrhein-Westfalen, Krefeld, 1992.

Kløve, B., Ala-Aho, P., Bertrand, G., Gurdak, J. J., Kupfersberger, H., Kværner, J., Muotka, T., Mykrä, H., Preda, E., Rossi, P., Uvo, C. B., Velasco, E., and Pulido-Velazquez, M.: Climate change impacts on groundwater and dependent ecosystems, J. Hydrol., in press, 10.1016/j.hydrol.2013.06.037, 2013.

Kooi, H.: Spatial variability in subsurface warming over the last three decades; insight from repeated borehole temperature measurements in The Netherlands. Earth Planet. Sci. Lett. 270, 86-94, 10.1016/g.epsl.2008.03.015.

Kurylyk, B. L., and MacQuarrie, K. T. B.: The uncertainty associated with estimating future groundwater recharge: A summary of recent research and an example from a small unconfined aquifer in a northern humid-continental climate, J. Hydrol., 492, 244-253, 2013.

Kurylyk, B. L., and MacQuarrie, K. T. B.: A new analytical solution for assessing climate change impacts on subsurface temperature, Hydrol. Processes, 28, 3161-3172, 10.1002/hyp.9861, 2014.

Kurylyk, B. L., P.-A. Bourque, C., and Macquarrie, K. T. B.: Potential surface temperature and shallow groundwater temperature response to climate change: An example from a small forested catchment in east-central NB (Canada), Hydrol. Earth Syst. Sci., 17, 2701-2716, 2013.

Kurylyk, B. L., MacQuarrie, K. T. B., and Voss, C. I.: Climate change impacts on the temperature and magnitude of groundwater discharge from shallow, unconfined aquifers, Water Resour. Res., 50, 3253-3274, 10.1002/2013WR014588, 2014a.

Kurylyk, B. L., McKenzie, J. E., MacQuarrie, K. T. B., and Voss, C. I.: Analytical solutions for benchmarking cold regions subsurface water flow and energy transport models: One-dimensional soil thaw with conduction and advection. Adv. Water. Res., 70, 172-184, 10.1016/j.advwatres.2015.05.005, 2014a.

Mann, M. E., and Schmidt, G. A.: Ground vs. surface air temperature trends: Implications for borehole surface temperature reconstructions, Geophys. Res. Lett., 30, 9-1, 2003.

Mann, M. E., Schmidt, G. A., Miller, S. K., and LeGrande, A. N.: Potential biases in inferring Holocene temperature trends from long-term borehole information, Geophys. Res. Lett., 36, 2009.

Mareschal, J. C., and Beltrami, H.: Evidence for recent warming from perturbed geothermal gradients: examples from eastern Canada, Clim. Dyn., 6, 135-143, 1992.

Marty, C.: Regime shift of snow days in Switzerland, Geophys. Res. Lett., 35, L12501, 10.1029/2008gl033998, 2008.

Maxwell, R. M., and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change, Nat. Geosci., 1, 665-669, 2008.

Mayer, T. D.: Controls of summer stream temperature in the Pacific Northwest, J. Hydrol., 475, 323-335, 2012.

Mellander, P. E., Löfvenius, M. O., and Laudon, H.: Climate change impact on snow and soil temperature in boreal Scots pine stands, Clim. Change, 85, 179-193, 2007.

Menberg, K., Steger, H., Zorn, R., Reuss, M., Proell, M., Bayer, P., and Blum, P.: Bestimmung der Wärmeleitfähigkeit im Untergrund durch Labor- und Feldversuche und anhand theoretischer Modelle [Determination of thermal conductivity in the subsurface using laboratory and field experiments and theoretical models], Grundwasser, 18, 103-116, 2013.

North, R. P., Livingstone, D. M., Hari, R. E., Köster, O., Niederhauser, P., and Kipfer, R.: The physical impact of the late 1980s climate regime shift on Swiss rivers and lakes, Inland Waters, 3, 341-350, 10.5268/IW-3.3.560, 2013.

Overland, J., Rodionov, S., Minobe, S., and Bond, N.: North Pacific regime shifts: Definitions, issues and recent transitions, Prog. Oceanogr., 77, 92-102, 10.1016/j.pocean.2008.03.016, 2008.

Pollack, H., Huang, S., and Shen, P.: Climate Change Record in Subsurface Temperatures: A Global Perspective, Science, 282, 279 - 281, 1998.

Reiter, M.: Possible Ambiguities in Subsurface Temperature Logs: Consideration of Ground-water Flow and Ground Surface Temperature Change, Pure Appl. Geophys., 162, 343-355, 10.1007/s00024-004-2604-4, 2005.

Risley, J. C., Constantz, J., Essaid, H., and Rounds, S.: Effects of upstream dams versus groundwater pumping on stream temperature under varying climate conditions, Water Resour. Res., 46, W06517, 10.1029/2009WR008587, 2010.

Rodionov, S. N.: A sequential algorithm for testing climate regime shifts, Geophys. Res. Lett., 31, L09204, 10.1029/2004gl019448, 2004.

Rodionov, S. N., and Overland, J. E.: Application of a sequential regime shift detection method to the Bering Sea ecosystem, ICES J. Mar. Sci., 62, 328-332, 10.1016/j.icesjms.2005.01.013, 2005.

Rodionov, S. N.: Use of prewhitening in climate regime shift detection, Geophys. Res. Lett., 33, L12707, 10.1029/2006gl025904, 2006.

Rudnick, D. L., and Davis, R. E.: Red noise and regime shifts, Deep Sea Res. Part I, 50, 691-699, 10.1016/S0967-0637(03)00053-0, 2003.

Saar, M. O.: Review: Geothermal heat as a tracer of large-scale groundwater flow and as a means to determine permeability fields, Hydrogeol. J., 19, 31-52, 10.1007/s10040-010-0657-2, 2011.

Schmidt, G. A., and Mann, M. E.: Reply to comment on "Ground vs. surface air temperature trends: Implications for borehole surface temperature reconstructions" by D. Chapman et al., Geophys. Res. Lett., 31, 2004.

Seidel, D. J., and Lanzante, J. R.: An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes, J. Geophys. Res., 109, D14108, 10.1029/2003jd004414, 2004.

Sharma, L., Greskowiak, J., Ray, C., Eckert, P., and Prommer, H.: Elucidating temperature effects on seasonal variations of biogeochemical turnover rates during riverbank filtration, J. Hydrol., 428-429, 104-115, 2012.

Smerdon, J. E., Pollack, H., Cermak, V., Enz, J. W., Kresl, M., Safanda, J., and Wehmiller, J. F.: Air-ground temperature coupling and subsurface propagation of annual temperature signals, J. Geophys. Res., 109, D21107 21101-21110, doi:10.1029/2004JD005056, 2004.

Stallman, R. W.: Steady one-dimensional fluid flow in a semi-infinite porous medium with sinusoidal surface temperature. J. Geophys. Res., 70, 2821-2827, 10.1029/JZ070i012p02821.

Stoll, S., Hendricks Franssen, H. J., Barthel, R., and Kinzelbach, W.: What can we learn from long-term groundwater data to improve climate change impact studies?, Hydrol. Earth Syst. Sci., 15, 3861-3875, 10.5194/hess-15-3861-2011, 2011.

Swanson, K. L., and Tsonis, A. A.: Has the climate recently shifted?, Geophys. Res. Lett., 36, L06711, 10.1029/2008gl037022, 2009.

Tague, C., Farrell, M., Grant, G., Lewis, S., and Rey, S.: Hydrogeologic controls on summer stream temperatures in the McKenzie River basin, Oregon, Hydrol. Processes, 21, 3288-3300, 10.1002/hyp.6538, 2007.

Taniguchi, M., Shimada, J., Tanaka, T., Kayane, I., Sakura, Y., Shimano, Y., Dapaah-Siakwan, S., and Kawashima, S.: Disturbances of temperature-depth profiles due to surface climate change and subsurface water flow: 1. An effect of linear increase in surface temperature caused by global warming and urbanization in the Tokyo metropolitan area, Japan, Water Resour. Res., 35, 1507-1517, 1999a.

Taniguchi, M., Williamson, D. R., and Peck, A. J.: Disturbances of temperature-depth profiles due to surface climate change and subsurface water flow: 2. An effect of step increase in surface temperature caused by forest clearing in southwest Western Australia, Water Resour. Res., 35, 1519-1529, 1999b.

Taylor, C. A., and Stefan, H. G.: Shallow groundwater temperature response to climate change and urbanization, J. Hydrol., 375, 601 - 612, 2009.

Taylor, R. G., Scanlon, B., Doll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J. F., Holman, I., and Treidel, H.: Ground water and climate change, Nat. Clim. Chang., 3, 322-329, 2013.

van Vliet, M. T. H., Ludwig, F., Zwolsman, J. J. G., Weedon, G. P., and Kabat, P.: Global river temperatures and sensitivity to atmospheric warming and changes in river flow, Water Resour. Res., 47, W02544, 10.1029/2010WR009198, 2011.

van Vliet, M. T. H., Franssen, W. H. P., Yearsley, J. R., Ludwig, F., Haddeland, I., Lettenmaier, D. P., and Kabat, P.: Global river discharge and water temperature under climate change, Global Environ. Chang., 23, 450-464, 10.1016/j.gloenvcha.2012.11.002, 2013.

Verein Deutscher Ingenieure: VDI 4640-1, Thermische Nutzung des Untergrundes: Grundlagen, Genehmigungen, Umweltaspekte (Thermal use of the underground: Fundamentals, approvals, environmental aspects), Beuth Verlag GmbH, Berlin, 2010.

von Storch, H.: Misuses of statistical analysis in climate research, in: Analysis of Climate Variability Applications of Statistical Techniques. , edited by: von Storch, H., and Navarra, A., Springer, New York, 1995.

Wenger, S. J., Isaak, D. J., Luce, C. H., Neville, H. M., Fausch, K. D., Dunham, J. B., Dauwalter, D. C., Young, M. K., Elsner, M. M., Rieman, B. E., Hamlet, A. F., and Williams, J. E.: Flow regime, temperature, and biotic interactions drive differential declines of trout species under climate change, Proc. Natl. Acad. Sci. U. S. A., 108, 14175-14180, 10.1073/pnas.1103097108, 2011.

Wirsing, G., and Luz, A.: Hydrogeologischer Bau und Aquifereigenschaften der Lockergesteine im Oberrheingraben (Baden-Württemberg). Freiburg i.Br., 2008.

Wu, H., Kimball, J. S., Elsner, M. M., Mantua, N., Adler, R. F., and Stanford, J.: Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers, Water Resour. Res., 48, W11530, 10.1029/2012WR012082, 2012.

Zhang, Y., Chen, W., Smith, S. L., Riseborough, D. W., and Cihlar, J.: Soil temperature in Canada during the twentieth century: Complex responses to atmospheric climate change, J. Geophys. Res., 110, D03112, 10.1029/2004JD004910, 2005.

Table 1: Location coordinates of the observation wells with basic information about the hydrological setting.

Well	Easting	Northing	Altitude [m asl]	Subsurface material	Distance to nearest stream
Dansweiler	2553462	5646975	88.2	fine to coarse sand, minor contents of gravel and silt[a]	~ 6 km (Erft)
Sinthern	2555310	5648820	64.4		~ 9 km (Erft)
Hardtwald 1	3457460	5435140	112.4	gravel and coarse sand with layers of fine sand and silt[b]	~ 6 km (Rhine)
Hardtwald 2	3457500	5435200	112.1		~ 6 km (Rhine)

[a] Klostermann, 1992, [b] HGK, 2007.

Table 2: Hydrogeological data of the four observation wells.

Well	Depth of water table [m bgl]	Depth of well screens [m bgl]	Average hydraulic conductivity [m s^{-1}]	Groundwater recharge rates [mm yr^{-1}]
Dansweiler	18 ± 1	22.5 - 22.6	1.0-5.0 × 10^{-4} [a]	221 ± 45[c]
Sinthern	16 ± 1	21.3 - 21.4	1.0-5.0 × 10^{-4} [a]	221 ± 45[c]
Hardtwald 1	7 ± 3	10 - 36	1.1-1.4 × 10^{-3} [b]	228 ± 45[d]
Hardtwald 2	7 ± 3	10.5 - 38.5	1.1-1.4 × 10^{-3} [b]	228 ± 45[d]

[a] Balke, 1973 [b] Wirsing and Luz, 2008 [c] Erftverband, 1995. [d] Deinlein, personal communication, 2013.

Table 3: Range of thermal properties and recharge values utilized in analytical solutions.

Location	Thermal conductivity (W m^{-1} K^{-1}) (min, mean, max)	Heat capacity (×10^{6} J m^{-3} °C^{-1}) (min, mean, max)	Thermal diffusivity (×10^{-7} m^{2} s^{-1}) (min, mean, max)	Recharge (mm yr^{-1}) (min, mean, max)
		Unsaturated zone		
Dansweiler	0.4, 1.2, 2.4	1.3, 2.0, 2.8	3.1, 5.8, 8.4	176, 221, 265
Sinthern	0.4, 1.1, 2.4	1.2, 2.0, 2.8	3.3, 5.6, 7.9	176, 221, 265
Hardtwald	0.8, 1.5, 2.4	1.5, 2.0, 2.5	5.1, 7.4, 9.6	182, 228, 274
		Saturated zone		
Dansweiler	1.5, 2.2, 3.1	2.4, 2.6, 2.8	6.2, 8.4, 10.9	176, 221, 265
Sinthern	1.4, 2.1, 3.0	2.3, 2.6, 2.8	6.1, 8.2, 10.7	176, 221, 265
Hardtwald	2.4, 2.9, 3.5	2.5, 2.7, 2.9	9.6, 10.8, 12.0	182, 228, 274

Table 4: Time lags and final p-values of the observed regime shifts in air and groundwater temperature. The specific years indicate the first year of the new regime. Time lags are defined as the period between the occurrence of a regime shift in local SAT and the corresponding successive shift in GWT.

Time-series	Regime shift late 1970s		Regime shift late 1980s			Regime shift late 1990s		
	Year	p-value	Year	Time lag to SAT (years)	p-value	Year	Time lag to SAT (years)	p-value
Global mean ΔT	1977	1.8×10^{-5}	1987	–	4.4×10^{-4}	1997	–	1.8×10^{-7}
Zonal mean ΔT	1977	9.1×10^{-4}	1988	–	4.6×10^{-3}	1997	–	5.7×10^{-5}
SAT Cologne	–	–	1988	–	6.7×10^{-4}	1999	–	5.5×10^{-3}
GWT Dansweiler	–	–	1991	+ 3 (±2)	1.1×10^{-9}	2003	+ 4 (±2)	1.0×10^{-1}
GWT Sinthern	–	–	1991	+ 3 (±2)	9.3×10^{-5}	2001	+ 2 (±2)	4.3×10^{-4}
SAT Karlsruhe	–	–	1988	–	3.2×10^{-5}	1999	–	3.8×10^{-2}
GWT Hardtwald 1	–	–	1989	+ 1 (±2)	3.6×10^{-4}	2000	+ 1 (±2)	1.3×10^{-4}
GWT Hardtwald 2	–	–	1989	+ 1 (±2)	9.2×10^{-4}	2000	+ 1 (±2)	1.0×10^{-2}

Table 5: Results (p-values) of the Mann-Kendall test for the absence of a trend for all regimes in SAT and GWT time-series.

Time-series	Regime I		Regime II		Regime III		Regime IV	
	Period	p-value	Period	p-value	Period	p-value	Period	p-value
Global mean ΔT	1950-1976	0.62	1977-1986	1.00	1987-1996	0.72	1997-2012	0.26
Zonal mean ΔT	1950-1976	0.30	1977-1986	0.64	1987-1996	0.47	1997-2012	0.72
SAT Cologne	–	–	1962-1987	0.40	1988-1998	0.89	1999-2011	0.46
SAT Karlsruhe	–	–	1962-1987	0.43	1988-1998	0.31	1999-2009	0.76
GWT Dansweiler	–	–	1970-1990	0.05	1991-2002	0.78	2003-2010	1.00
GWT Sinthern	–	–	1974-1990	0.06	1991-2000	0.18	2001-2006	0.01
GWT Hardtwald 1	–	–	1968-1988	0.22	1989-1999	0.14	2000-2011	0.13
GWT Hardtwald 2	–	–	1968-1988	0.43	1989-1999	0.59	2000-2010	0.31

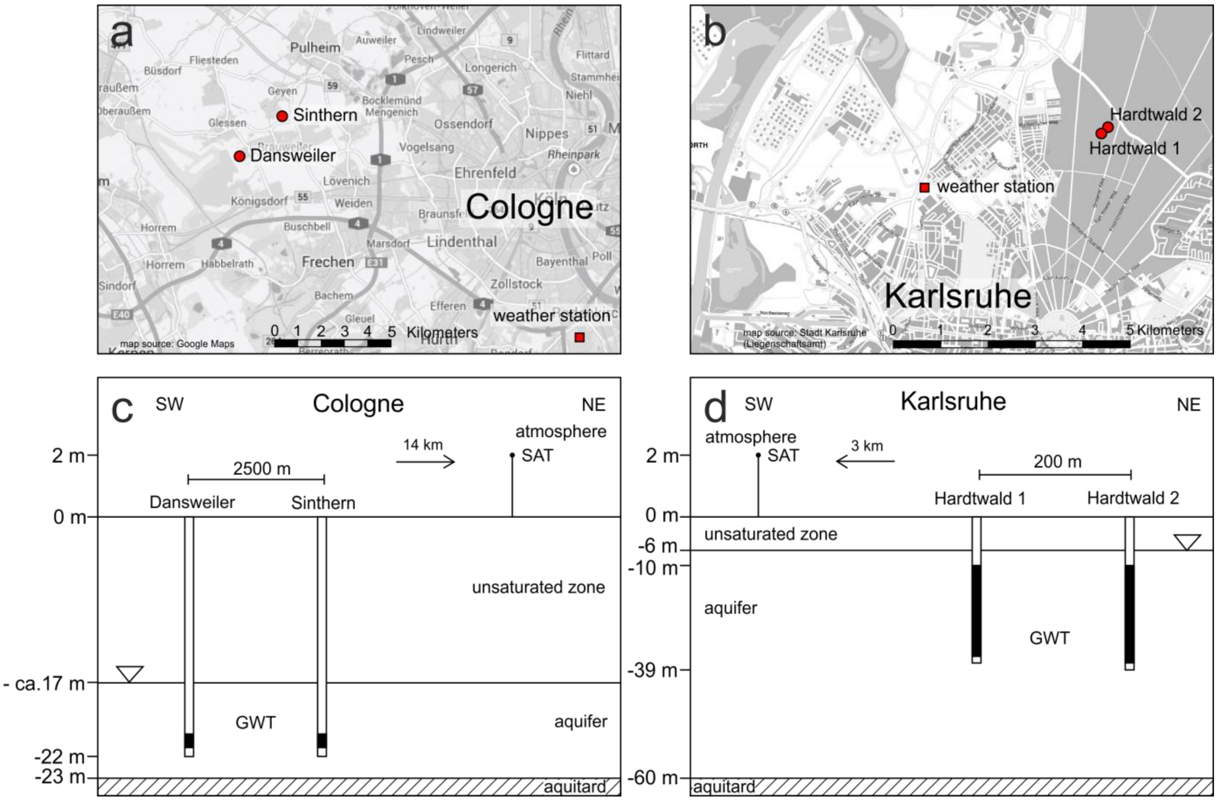

Figure 1: a, b: Locations of the four observation wells and two weather stations used in the present study. c, d: Conceptual sketch of the well settings in the aquifers close to Cologne (left) and Karlsruhe (right). The black zones in the wells indicate the location of the filter screens. Please note the different scales in the subsurface.

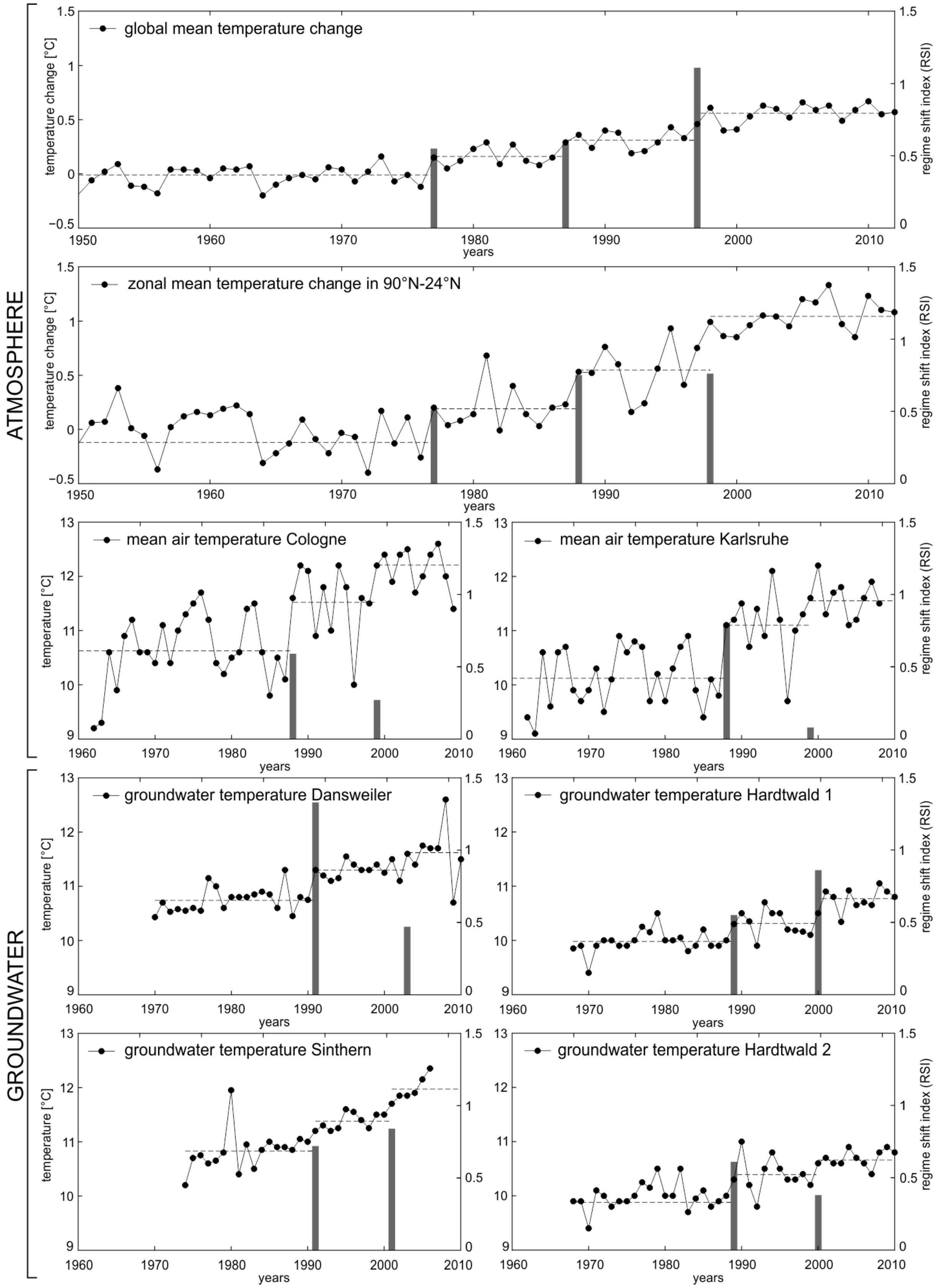

Figure 2: Time series of temperature data with long-term means (dashed lines) and observed regime shift with RSI values.

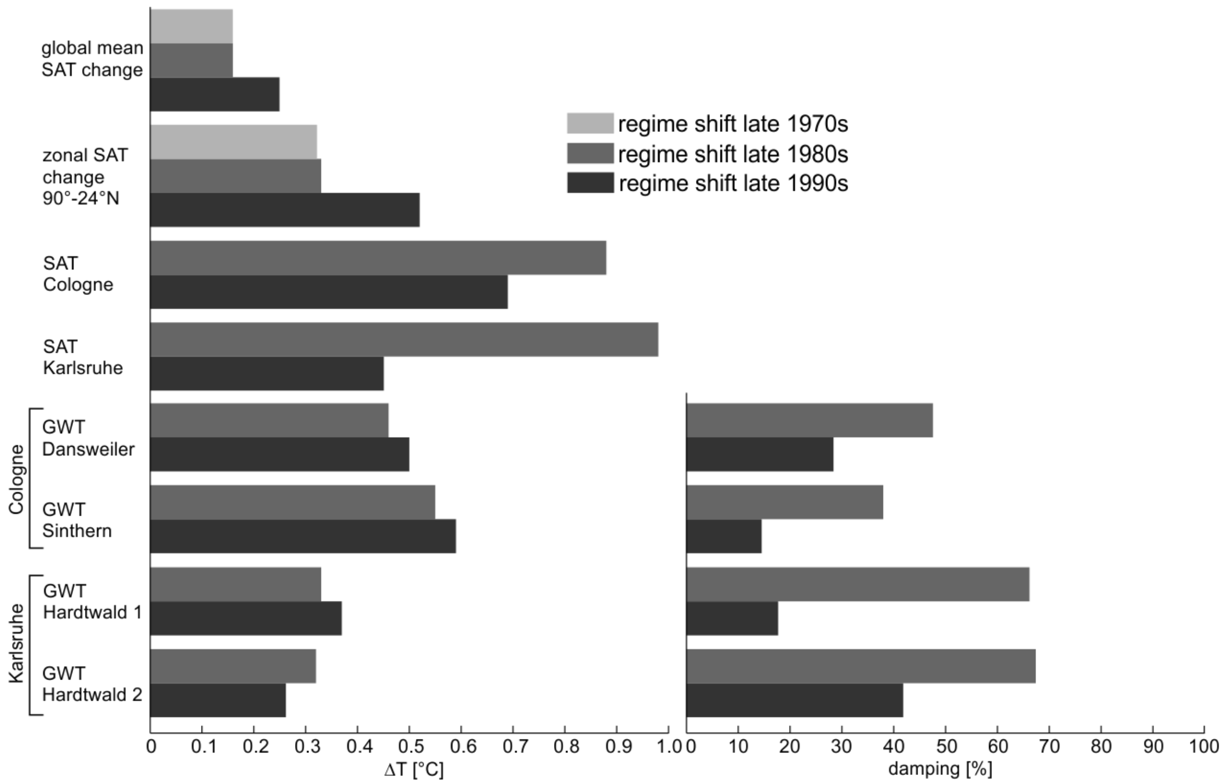

Figure 3: Left: Magnitude of regime shifts in time-series of atmospheric and groundwater temperatures. Right: relative damping of the regime shift magnitude in groundwater temperatures compared to regional atmospheric regime shift, calculated as 100 minus the ratio of ΔT in GWT to ΔT in SAT in percent.

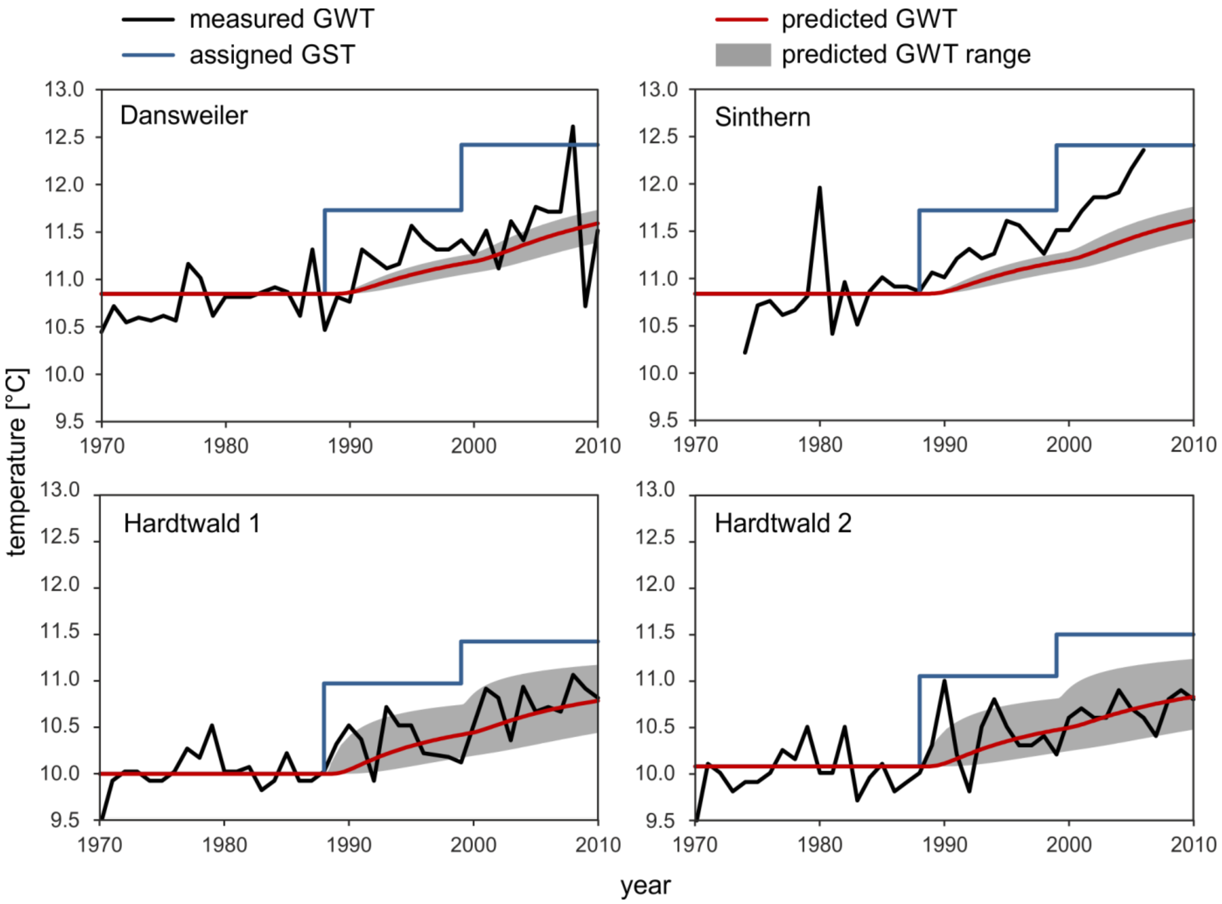

Figure 4: Measured GWT, predicted GWT, and assigned GST boundary conditions for the TCA model (Eq. 6) for each well versus the year. Red lines indicate GWT results obtained using the mean well screen depth, thermal properties and recharge rates presented in Table 2 and 3. The GWT data at the lower and higher ends of the temperature envelope are obtained with the ranges in well screen depth, thermal diffusivity, heat capacity, and recharge rates (Table 2 and 3).